大数据背景下新闻编辑的数据挖掘创新素能研究

摘　要：信息时代数据资源的快速膨胀与传播，对新闻编辑来说，每天要接触和阅读大量的信息，这些庞杂的信息如何进行有效的提炼，往往考验着新闻编辑的数据挖掘创新能力。新闻报道在大数据时代下，新闻生产与呈现方式的变化，势必对新闻从业者提出更高挑战。新闻编辑作为新闻报道的主体，更需要从知识结构、数据量爆炸式增长中提升数据的搜集、加工、呈现水平，才能更好地适应媒介新环境的变化，为受众提供有吸引力、说服力和影响力的新闻信息。

关键词：大数据；新闻编辑；数据挖掘；创新素能
中图分类号：G214　　文献标识码：A
文章编号：1671-0134（2017）07-035-02　　DOI：10.19483/j.cnki.11-4653/n.2017.07.003

■文 / 侯　丽

从网络新闻到自媒体的全面覆盖，学界在面对爆炸式增长的新闻类数据资源时，将更多的视线聚焦在“大数据”的挖掘上。大数据已经成为当前新闻编辑工作者需要具备的独特素能，要能够从众多的新闻数据中搜集和提炼新闻点，并将之与现代媒介新闻传播生产方式相融合。数据新闻作为业界探讨的热点议题，源自“数据新闻学”概念，并由此延伸至新闻传播实践中。新闻编辑是对新闻信息进行编辑加工的人，以新闻职业嗅觉来参与到新闻媒介形态的变化中，而随着大数据新闻信息采集的复杂性，新闻编辑的数据挖掘素能就变得尤为重要。

1. 大数据下新闻编辑工作面临的挑战

在《数据新闻手册》中对“数据新闻学”的解释为“基于数据的抓取、挖掘、统计、分析和可视化呈现的新型新闻报道方式，其核心就是用数据报道新闻”。事实上，早在 2012 年《纽约时报》将“大数据”解释为“基于数据和分析做出的决策”，而非“基于经验和直觉”。从近年来各国新闻主流媒体对数据新闻学的研究与实践来看，业界和学界都提出了探索与尝试。比如我国的新浪网“图解天下”栏目、网易的“数读”栏目，搜狐的“数字之道”等，都是建立在数据新闻报道基础上。作为新闻编辑工作者，在面对大数据时代数据新闻学的挑战中，如何快速发现信息的价值？如何挖掘有意义的信息？如何从海量的数据资源中，利用不同的工具来分析和梳理新闻事实？重构新闻编辑的数据挖掘创新能力就显得尤为重要。一是受众对新闻信息多样化需求，对来自电视、报纸、网络的各类信息资源，迫切需要从尽可能少的时间里来获得“碎片化阅读”“快餐式阅读”。美国有学者机构对“读者如何读报”进行调研发现，其中 3/4 的读者关注的是图片和图示；1/2 的读者关注的是标题；1/3 的读者关注的是提要和图片说明；1/4 的读者关注的是文字报道。可见，从读者的阅读诉求上，图片、图示、图表等表现形式更具有吸引力，新闻编辑应该有意识地从营造“轻量化阅读体验”中，尽可能地融入易懂、易看、易理解的图示化信息。二是新闻编辑自身在面对大数据时，要能够化繁为简、化抽象为具体，简言之，就是要能够从丰富的媒介信息资源中挖掘有价值的、有意义的新闻信息及表现形式。作为面向社会舆论的新闻编辑，首先要能够从复杂的新闻事实中，讲述好新闻问题，对整个新闻事件进行清晰的、动态的展示。学者杰夫 · 麦吉在探讨新闻的媒介责任时提出“新闻是面向公众解释复杂难懂数据的过程，新闻编辑的任务就是从数据中来提供易于理解、具有说服力的数据新闻”。因此，新闻编辑的数据挖掘创新能力，需要从自我知识结构、数据挖掘能力中来提供有价值的新闻信息，呈现给媒介受众。

2. 大数据下新闻编辑的数据挖掘创新素能

数据新闻挖掘的源泉是数据，而数据挖掘创新能力需要从搜集、加工、呈现三个阶段来完成。

2. 1 新闻编辑的数据搜集素能

数据本身是多元的，数据新闻是以公开为基础的各类新闻数据，而新闻编辑在面对新闻数据时，又很容易受到海量数据的影响而迷失方向。如何开展数据搜集，需要从三种渠道来实现。一是利用网络搜索引擎，比如百度、谷歌等搜索工具；二是利用专业数据库，比如一些新闻类数据中心数据库；三是网络论坛及其他网络新闻数据站点。对于搜索引擎的使用是最常见也是最广泛的一种方式，很多新闻编辑都会使用，但对于新闻搜索实践中的“关键词”选择，往往考查着新闻编辑的新闻嗅觉力。另外，在搜索引擎高级检索功能的使用，比如附加 · mdb 搜索 Access 数据库，利用 · xls 来搜索相关的数据表格等。借助于网络论坛、网络数据站点来搜集新闻数据，或者通过网络调查网站及网络投票平台来获取数据，都是建立在新闻调查问卷基础上，以某些主题的论

坛为表现形式，从而获得最新的数据资源。新闻编辑在搜集数据时，渠道并非仅限于此，还可以拓宽到其他各类专业数据库。

2.2 新闻编辑的数据加工素能

搜集数据只是新闻编辑获取更多新闻资源的基础和前提，如何将这些分散的、独立的、隐匿的数据进行整合，并从中提炼出有价值的新闻数据信息，就需要新闻编辑具有相应的数据加工素能。数据加工素能本质上是从庞杂的数据资源中来提炼信息点，为受众提供有价值的新闻内容，让新闻传播更具有可读性。面对海量新闻数据，新闻编辑要善于应用数据整合方法。比如对一些不必要大数据进行剔除，保留有效的精要数据，而且是准确无误的；对相关数据的格式进行统一，特别是对于一些数据库表信息进行有效转换，使其具备可读性、完备性；对于一些抽象的数据资源，还需要通过换算、对比等方法来活用。比如一些统计资料中的数据，需要从相关数据库的对比性中来条理化，避免数据的冗杂、重复。加工数据的过程是新闻编辑需要谨慎而严肃对待的，特别是一些概念性数据、逻辑性数据，要避免出现混杂或错误。对于这些报道中的编辑加工问题，也是常见的错误，有失科学性。通常情况下，在新闻报道中，引入数据时往往是对新闻核心内容的评述，新闻编辑在数据加工时，更应该从严肃、认真、准确、客观上来把好关，要提高职业警惕性，要减少和消除各类数据引用错误问题，保障新闻报道的准确性、科学性。对于新闻数据的加工环节，新闻编辑要对一些引用的数据，特别是一些数字进行严格审查，增强数据敏感性。

2.3 新闻编辑的数据呈现素能

新闻报道的最终结果是为受众呈现可读性强的新闻稿件，以及可视化新闻报道等内容。这些新闻信息，所涉及的技术领域较为宽泛，有图形类数据呈现方式，有计算表格类数据等。面对大数据时代下不同受众对新闻获取渠道的多样性，在呈现新闻数据可视化过程中，要突出新闻信息数据、新闻动态图表的有序呈现。比如新闻编辑要能够利用可视化技术，将抽象的、复杂的数据转换为生动的、趣味性的新闻报道。在一些可视化技术应用中，新闻编辑要不断提升自我的技术能力。比如在 ManyEyes 网站，只要上传相关数据，就可以获得可视化图表；当然，在呈现图表方式上，ManyEyes 网站还提供多种可视化方案，如 Word 格式、TreeMap 层次格式，以及可视化人物关系网络图等。再者，谷歌也推出了可视化数据服务功能，可以实现对相关数据的批量处理，如筛选、分类、合并、聚合等。这些免费的数据呈现工具和软件，为新闻编辑优化新闻数据呈现方式提供了便利。

3. 结语

当然，新闻编辑在面对数据新闻学的发展中，还要从自我知识结构、编辑处理技能学习和培训中，增强自我对新闻的敏锐力、挖掘力、表现力。比如学习一些艺术设计软件技巧，增强自我新闻的版面设计效果，注重色彩的搭配，调整不同新闻内容的均衡性，让自己的新闻更具有魅力和张力。媒

参考文献

[1] 韩新明，刘海明．数据新闻编辑应该注意的几个问题 [J]. 中国出版，2016（09）：34-36.

[2] 王亚雷．“微媒介”环境下传统新闻编辑的转型发展之路 [J]. 采写编，2017（01）：55.

[3] 石勤．善用数据新闻 创新报道方式 [J]. 中国广播，2015（02）：63-65.

（作者单位：承德广播电视台）

（上接第 34 页）

10.2139/ssrn.2094215.

[4] Bleicher，A.（2014）. 5G service on your 4G phone? [News]. IEEE Spectrum，51（4），13-18.

[5] Bhowmik，A. K.（2016）. 35 - 1：Invited Paper：Real - Time 3D - Sensing Technologies and Applications in Interactive and Immersive Devices. Sid Symposium Digest of Technical Papers，47（1），440-443.

[6] Russell，S.，& Norvig，P.（1995）. Artificial intelligence：a modern approach. Artificial Intelligence，，369-380.

[7][美] 保罗・莱文森著，何道宽译．新新媒介．复旦大学出版社，2014（2）.

[8] Kwak，H.，Lee，C.，Park，H.，& Moon，S.（2010）. What is Twitter，a social network or a news media?. international world wide web conferences，591-600.

[9][美] 尼葛洛庞帝著，胡冰译．数字化生存．海南出版社，1996（1）.

[10] Starostin，E. L.，& Gh，V. D. H.（2007）. The shape of a möbius strip. Nature Material，6（8），563-567.

（作者单位：华中师范大学新闻传播学院）